Maximum entropy moment system of the semiconductor Boltzmann equation using Kane's dispersion relation

Michael Junk* Vittorio Romano†

Abstract

It is known that the maximum entropy moment systems of the gas-dynamical Boltzmann equation suffer from severe disadvantages which are related to the non-solvability of an underlying maximum entropy moment problem unless restrictions on the choice of the macroscopic variables are made. In this article, we show that no such difficulties appear in the semiconductor case if Kane's dispersion relation is used for the energy band of electrons.

Keywords. maximum entropy moment closure, semiconductor Boltzmann equation, Kane's dispersion relation

1 Introduction

The direct integration of the transport equation coupled to the Poisson equation for the description of the motion of charges in semiconductors is a daunting computational task. Since one usually is not interested to the complete details of the distribution function but to quantities as average electron density, energy, velocity, several macroscopic models for charge transport in semiconductors have been developed. For a complete review on the subject the interested reader can see [1].

These models are based on the moment systems arising from the Boltzmann equation and require suitable closure assumptions. A physically sound way to get the sought closure relations is based on the maximum entropy principle (hereafter MEP). It is based on the information theory of Shannon [2] and has been introduced in statistical physics in [3, 4]. From a mathematical point of view MEP gives an approximation of the exact distribution function in terms of a finite number of moments and require to solve a constrained optimization problem. However this latter does not always admit a solution. This disadvantage can been seen in the case of gas dynamics when one considers moments with respect to weight functions represented by polynomials in the microscopic velocity of degree higher than two [5, 6, 7, 8].

Since the parabolic approximation for the energy band of electrons leads to a moment system with the same type of weight functions, the same drawback of the gas dynamics arises.

*FB Mathematik, Universität Kaiserslautern, Erwin-Schrödingerstraße, 67663 Kaiserslautern, Germany,(junk@mathematik.uni-kl.de)

†Dipartimento di Matematica e Informatica, Università di Catania, viale A. Doria 6 - 95125 Catania, Italy,(romano@dmi.unict.it)

The main goal of the present article is to show that such a problem is overcome when one employs the Kane model for the energy band. It is proved that the corresponding maximum entropy models are symmetric hyperbolic systems with convex domains of definition and that the equilibria are interior points, guaranteeing the validity of expansions around equilibrium states.
The plan of the article is the following. In section 2 the maximum entropy problem is presented for the moment systems arising in charge transport inside semiconductors. The solvability of such a problem is analyzed in section 3 while in the last section the cases of the Euler-Poisson model and the 8-moment one are treated in more detail, giving numerical results about the realizability region.

2 The maximum entropy moment systems for electrons in semiconductors

We consider a Boltzmann type model to describe the dynamics of electrons in semiconductor lattices. In a semi classical approximation [9], a kinetic description of electrons in a semiconductor is given by a transport equation for the one particle distribution function $f(t, \boldsymbol{x}, \boldsymbol{k})$, which represents the probability of finding an electron at time t in an elementary volume $d\boldsymbol{x}d\boldsymbol{k}$, around position $\boldsymbol{x}$ and with crystal momentum $\boldsymbol{k}$,

$$\frac{\partial f}{\partial t} + v_i(\boldsymbol{k})\frac{\partial f}{\partial x_i} - \frac{e}{\hbar}E_i\frac{\partial f}{\partial k_i} = \mathcal{C}[f]. \tag{1}$$

Here e is the absolute value of the electron charge, $\boldsymbol{k}$ represents the crystal momentum of the electron and $\boldsymbol{E}$ is the electric field which is related to the electron distribution by Poisson's equation:

$$\boldsymbol{E} = -\nabla\phi, \qquad \epsilon\Delta\phi = -e(N_D - N_A - n),$$

where ϕ is the electric potential and ϵ the permittivity of the semiconductor. N_D and N_A are respectively the donor and acceptor density. They depend only on $\boldsymbol{x}$ and are considered as known functions. n is the electron density which is related to f by

$$n = \int_B f d\boldsymbol{k},$$

B being the first Brillouin zone. This latter is a set of positive Lebesgue measure, symmetric with respect to the origin. Its properties can be found in the textbooks of solid state physics, e.g. in [10]. The right hand side $\mathcal{C}[f]$ in (1) is the collision operator, which takes into account scattering of the electrons with acoustical and optical phonons and with impurities (for further details see [11, 12]). The electron velocity $\boldsymbol{v}(\boldsymbol{k})$ depends on the electron energy $\mathcal{E}$ by the relation

$$\boldsymbol{v}(\boldsymbol{k}) = \frac{1}{\hbar}\nabla_{\boldsymbol{k}}\mathcal{E}.$$

In general, the expression of $\mathcal{E}$ (the so called band structure) depends on the material and is very complicated. A simple approximation is given by the *parabolic band*: the effective mass is a constant scalar m^*, the relation between energy and wave vector is

$$\mathcal{E}(\boldsymbol{k}) = \frac{\hbar^2|\boldsymbol{k}|^2}{2m^*}, \qquad \boldsymbol{k} \in \mathbb{R}^3,$$

and therefore

$$\boldsymbol{v}(\boldsymbol{k}) = \frac{\hbar}{m^*}\boldsymbol{k}.$$

A more refined approximation is *Kane's dispersion relation* which takes into account the non-parabolicity at high energies. In this case

$$\mathcal{E}(\boldsymbol{k}) = \frac{1}{1+\sqrt{1+2\frac{\alpha}{m^*}\hbar^2|\boldsymbol{k}|^2}}\frac{\hbar^2|\boldsymbol{k}|^2}{m^*} = \sqrt{\frac{1}{4\alpha^2}+\frac{\hbar^2|\boldsymbol{k}|^2}{2\alpha m^*}} - \frac{1}{2\alpha}, \qquad \boldsymbol{k} \in \mathbb{R}^3$$

where $\alpha > 0$ is the non-parabolicity parameter. The corresponding electron velocity is

$$\boldsymbol{v}(\boldsymbol{k}) = \frac{1}{\sqrt{1+\frac{2\alpha}{m^*}\hbar^2|\boldsymbol{k}|^2}}\frac{\hbar}{m^*}\boldsymbol{k}.$$

Besides the electron density n, other physically relevant quantities are the average electron velocity $\boldsymbol{u}$ relative to the crystal, assumed to be at rest,

$$\boldsymbol{u} = \frac{1}{n}\int_{\mathbb{R}^3} \boldsymbol{v}(\boldsymbol{k}) f \, d\boldsymbol{k},$$

the average electron energy

$$W = \frac{1}{n}\int_{\mathbb{R}^3} \mathcal{E}(\boldsymbol{k}) f \, d\boldsymbol{k},$$

and the flux of energy

$$\boldsymbol{S} = \frac{1}{n}\int_{\mathbb{R}^3} \boldsymbol{v}(\boldsymbol{k})\mathcal{E}(\boldsymbol{k}) f \, d\boldsymbol{k}.$$

In other words, all the quantities of interest can be written as suitable moments of the distribution function f. To generalize this observation, we introduce general weight functions $a_i : \mathbb{R}^d \mapsto \mathbb{R}$ and the corresponding moments

$$\rho_i = \langle f, a_i \rangle, \qquad i = 1, \ldots, m$$

where $\langle \cdot, \cdot \rangle$ denotes $\boldsymbol{k}$ integration. Of course in the case of physical interest $d = 3$. Observing that both $n, \boldsymbol{u}$, and $W, \boldsymbol{S}$ are $\boldsymbol{v}$-polynomial moments of mass density f and energy density $\mathcal{E}f$ respectively, we split the vector of weight functions $\boldsymbol{a}$ into two subgroups. The first m_1 components of $\boldsymbol{a}$ are chosen as $(P_1(\boldsymbol{v}(\boldsymbol{k})), \ldots, P_{m_1}(\boldsymbol{v}(\boldsymbol{k})))$ where $P_1, \ldots, P_{m_1}$ are linearly independent polynomials with $P_1(\boldsymbol{v}) = 1$, and the remaining m_2 components give rise to energy moments $(\mathcal{E}(\boldsymbol{k})Q_1(\boldsymbol{v}(\boldsymbol{k})), \ldots, \mathcal{E}(\boldsymbol{k})Q_{m_2}(\boldsymbol{v}(\boldsymbol{k})))$ where, again, $Q_1, \ldots, Q_{m_2}$ are linearly independent polynomials and $Q_1(\boldsymbol{v}) = 1$.
We remark that the choice of the weight functions leads to a system of balance law which does not satisfy Galilean invariance. In fact, as shown in [13], Galilean invariance requires that the weight functions are polynomials in $\boldsymbol{v}$. As a consequence, the present approach does not help to cure the problem of the maximum entropy approach in the gas dynamical case. In the case of semiconductor equations, however, the non-Galilean invariance is part of the model because the equations are written in a reference frame comoving with the semiconductor crystal (see [14]).
Since the direct numerical approximation of the kinetic equation (1) is very expensive due to the high dimensionality of the problem, and in view of the fact that one is rather

interested in moments of f than in f itself, it is a natural idea to derive equations directly for the averaged quantities. Multiplying (1) with weight functions $\boldsymbol{a} = (a_1, \dots, a_m)^T$ and integrating over $\boldsymbol{k}$ (abbreviated by $\langle \cdot, \cdot \rangle$), we obtain equations for the moments

$$\frac{\partial \boldsymbol{\rho}}{\partial t} + \frac{\partial}{\partial x_j} \langle f, v_j \boldsymbol{a} \rangle = \langle \mathcal{C}[f] + \gamma \boldsymbol{E} \cdot \nabla_{\boldsymbol{k}} f, \boldsymbol{a} \rangle , \qquad \gamma = e/\hbar. \tag{2}$$

The system would be closed if the particle distribution could be expressed in terms of the moment vector $\boldsymbol{\rho}$

$$f(t, \boldsymbol{x}, \boldsymbol{k}) = F(\boldsymbol{\rho}(t, \boldsymbol{x}), \boldsymbol{k})$$

A method to obtain such a relationship is the maximum entropy approach where $F(\boldsymbol{\rho}, \boldsymbol{k})$ is taken as solution of the problem

$$\begin{gathered} \text{maximize } H(f) = -\langle f, \log f - 1 \rangle \\ \text{with } f \geq 0 \text{ and } \langle f, \boldsymbol{a} \rangle = \boldsymbol{\rho} \end{gathered} \tag{3}$$

Variants and generalizations of this basic idea have been pursued by several authors [4, 5, 15, 16, 17, 18] and in [1, 14, 19, 20, 21, 22] the approach has been applied to the semiconductor Boltzmann equation.

For general a_i, the formal solution of (3) is obtained with the method of Lagrange multipliers. We introduce the Lagrange functional

$$L(f, \boldsymbol{\lambda}) := H(f) - \boldsymbol{\lambda} \cdot (\boldsymbol{\rho} - \langle f, \boldsymbol{a} \rangle)$$

where $\boldsymbol{\lambda}$ is the vector of Lagrange multipliers. The necessary condition that all directional derivatives vanish in the maximum $f_{\boldsymbol{\lambda}}$ leads to

$$0 = \delta L(f_{\boldsymbol{\lambda}}, \boldsymbol{\lambda}) = (-\log f_{\boldsymbol{\lambda}} + \boldsymbol{\lambda} \cdot \boldsymbol{a})\, \delta f_{\boldsymbol{\lambda}}$$

so that

$$f_{\boldsymbol{\lambda}} = \exp(\boldsymbol{\lambda} \cdot \boldsymbol{a}). \tag{4}$$

Finally, the Lagrange multipliers $\boldsymbol{\lambda}$ are chosen in such a way (if possible) that the moment constraints $\boldsymbol{\rho} = \langle f_{\boldsymbol{\lambda}}, \boldsymbol{a} \rangle$ are satisfied which gives rise to a function $\boldsymbol{\lambda} = \boldsymbol{\lambda}(\boldsymbol{\rho})$. We then introduce $F(\boldsymbol{\rho}, \boldsymbol{k}) = f_{\boldsymbol{\lambda}(\boldsymbol{\rho})}(\boldsymbol{k})$.

Using the maximum entropy distribution, we can now close the moment system (2) and obtain

$$\frac{\partial \boldsymbol{\rho}}{\partial t} + \frac{\partial}{\partial x_j} \boldsymbol{G}_j(\boldsymbol{\rho}) = \boldsymbol{P}(\boldsymbol{\rho}) \tag{5}$$

where $\boldsymbol{G}_j$ and $\boldsymbol{P}$ are given by

$$\boldsymbol{G}_j(\boldsymbol{\rho}) = \langle F(\boldsymbol{\rho}), v_j \boldsymbol{a} \rangle , \quad \boldsymbol{P}(\boldsymbol{\rho}) = \langle \mathcal{C}[F(\boldsymbol{\rho})] + \gamma \boldsymbol{E} \cdot \nabla_{\boldsymbol{k}} F(\boldsymbol{\rho}), \boldsymbol{a} \rangle$$

It can be shown (see, for example, [5, 17]) that $\eta(\boldsymbol{\rho}) = -H(F(\boldsymbol{\rho}))$ is a (locally) strictly convex entropy for the system (5). Moreover, from the properties of the collision operator, the positivity of the entropy production has been proved in [25, 26, 27]. The existence of a convex entropy implies that (5) is symmetric hyperbolic. However, this nice property alone does not guarantee practicability of the model (5). Depending on the choice of weight functions a_i, it can happen that problem (3) is not always solvable, i.e. that there exist

moment vectors $\boldsymbol{\rho}$ which cannot be written as $\boldsymbol{a}$-moments of any exponential density $f_{\boldsymbol{\lambda}} = \exp(\boldsymbol{\lambda}\cdot\boldsymbol{a})$. Since the domain of definition $\mathcal{U}$ of $\boldsymbol{G}_j$ and $\boldsymbol{P}$ is given by those moment vectors for which the solution of (3) exists, the non-solvability implies that $\mathcal{U}$ does not coincide with the set of all $\boldsymbol{a}$-moments (which is an open, convex cone). This structural deficiency of $\mathcal{U}$ is accompanied by the disadvantage that the equilibrium states are located on $\partial\mathcal{U}$ and that they are singular points of the flux functions $\boldsymbol{G}_j$. This has been demonstrated for maximum entropy moment systems which are based on polynomial weight functions [6, 7, 23, 24]. Since the parabolic band approximation also leads to such polynomial weights, similar conclusions apply.
Our main goal in this article is to show that Kane's model is superior to the parabolic band approximation in the sense that the corresponding moment system has a nice mathematical structure: it is a symmetric hyperbolic system with an open and convex domain of definition. The equilibria are interior points and the fluxes are regular at these states so that expansions around equilibria are reasonable in contrast to the parabolic band case.
As already mentioned, the hyperbolicity is an immediate structural feature of the moment system and since equilibria are contained in $\mathcal{U}$, they have to be interior points if the domain of definition is open. The smoothness of the fluxes follows from the inverse function theorem using the fact that $\boldsymbol{\lambda} \mapsto \langle f_{\boldsymbol{\lambda}}, \boldsymbol{a}\rangle$ is continuously differentiable with a positive definite Jacobian matrix $\langle f_{\boldsymbol{\lambda}}, \boldsymbol{a}\otimes\boldsymbol{a}\rangle$: for any vector $0 \neq \boldsymbol{\xi} \in \mathbb{R}^m$, we have

$$\sum_{i,j=1}^{m} \langle f_{\boldsymbol{\lambda}}, a_i a_j\rangle\, \xi_i\xi_j = \left\langle f_{\boldsymbol{\lambda}}, \left(\sum_{i=1}^{m} \xi_i a_i\right)^2 \right\rangle$$

which is strictly positive if the weight functions are, for example, continuous and linearly independent. Thus, what remains to be checked is that $\mathcal{U}$ is open and convex. We prove this fact by showing the solvability of (3) for all possible moment vectors $\boldsymbol{\rho}$, or in other words, by showing that $\mathcal{U}$ coincides with the open convex cone of all $\boldsymbol{a}$-moments.

3 Solvability of the maximum entropy problem

3.1 Statement of the main result

In order to state our main result, we first reformulate (3). For notational convenience, we measure $\mathcal{E}$, $\boldsymbol{k}, \boldsymbol{v}$ in units $1/(2\alpha)$, $\sqrt{m^*/(2\alpha\hbar^2)}$, and $1/\sqrt{2\alpha m^*}$ which leads to

$$\mathcal{E}(\boldsymbol{k}) = \sqrt{1+|\boldsymbol{k}|^2} - 1, \qquad \boldsymbol{v}(\boldsymbol{k}) = \frac{\boldsymbol{k}}{\sqrt{1+|\boldsymbol{k}|^2}}. \tag{6}$$

For small $\boldsymbol{k}$, we see a similarity to the parabolic band approximation because $\mathcal{E}(\boldsymbol{k}) \sim |\boldsymbol{k}|^2/2$ and $\boldsymbol{v}(\boldsymbol{k}) \sim \boldsymbol{k}$. For large $\boldsymbol{k}$, however, $\boldsymbol{v}(\boldsymbol{k})$ is bounded and $\mathcal{E}(\boldsymbol{k})$ grows only linearly due to the estimates

$$|\boldsymbol{v}(\boldsymbol{k})| < 1, \qquad |\boldsymbol{k}| - 1 \leq \mathcal{E}(\boldsymbol{k}) \leq 2|\boldsymbol{k}| + 1. \tag{7}$$

Based on $\mathcal{E}$ and $\boldsymbol{v}$ and two sets $\{P_1, \ldots, P_{m_1}\}$, $\{Q_1, \ldots, Q_{m_2}\}$ of linearly independent polynomials with $P_1 = Q_1 = 1$, we define the weight functions as

$$\boldsymbol{a} = (P_1(\boldsymbol{v}), \ldots, P_{m_1}(\boldsymbol{v}), \mathcal{E}Q_1(\boldsymbol{v}), \ldots, \mathcal{E}Q_{m_2}(\boldsymbol{v}))^T. \tag{8}$$

Since the assumption of a three dimensional $\boldsymbol{k}$-space is not relevant for our argument, we assume $\boldsymbol{k} \in \mathbb{R}^d$. The moment set related to the weights a_i is generated by the functions in

$$\mathcal{F} = \{f \geq 0 \,:\, f \not\equiv 0,\, |\boldsymbol{a}|f \in \mathbb{L}^1(\mathbb{R}^d)\}. \tag{9}$$

Here, $f \geq 0$ and $f \not\equiv 0$ are to be understood in the measure theoretic sense, i.e. $\{\boldsymbol{x} : f(\boldsymbol{x}) > 0\}$ should have positive Lebesgue measure. The corresponding moments are collected in

$$\mathcal{M} = \{\langle f, \boldsymbol{a}\rangle : f \in \mathcal{F}\}. \tag{10}$$

Using this notation and the definition of the entropy functional

$$H(f) = -\langle f, \log f - 1\rangle\,, \tag{11}$$

we can restate (3) as

$$\begin{gathered} \text{maximize } H(f) \\ \text{subject to } f \in \mathcal{F} \text{ and } \langle f, \boldsymbol{a}\rangle = \boldsymbol{\rho} \end{gathered} \tag{12}$$

Our main result is

Theorem 1 *The maximum entropy moment problem* (12) *is uniquely solvable for any* $\boldsymbol{\rho}$ *inside the open, convex cone* $\mathcal{M}$*. The solution is an exponential density* $\exp(\boldsymbol{\lambda} \cdot \boldsymbol{a})$ *for some* $\boldsymbol{\lambda} \in \mathbb{R}^m$ *depending on* $\boldsymbol{\rho}$*.*

As already mentioned, a similar result does not hold for the maximum entropy moment problem with polynomial weight functions arising in connection with the parabolic band approximation. In this case, one can find moment vectors in $\mathcal{M}$ which are arbitrarily close to the moment vector of a Maxwellian but for which the maximum entropy problem is *not* solvable. This may seem surprising in view of Theorem 1 and the fact that, for small $\boldsymbol{k}$, the moment functions in Kane's approach essentially coincide with polynomial moments. However, the non-solvability is a consequence of the behavior of the exponential densities for large $\boldsymbol{k}$ which is quite different for the two models.
In order to prove Theorem 1, we are going to use a general result of Csiszar about the solvability of minimum relative entropy problems on sets of probability measures [28]. The connection between the two results is based on a few simple transformations. First, we observe that, up to normalization, every $f \in \mathcal{F}$ can be viewed as a probability density. The normalization $f^* = f/\langle f, 1\rangle$ is abbreviated by a $*$-superscript and its image of $\mathcal{F}$ is denoted $\mathcal{F}^*$. Since we assume $a_1 = 1$, the moment vector of f^* has the structure

$$\langle f^*, \boldsymbol{a}\rangle = (1, \rho_2/\rho_1, \ldots, \rho_m/\rho_1)^T, \qquad \boldsymbol{\rho} = \langle f, \boldsymbol{a}\rangle\,,$$

which gives rise to a normalization operation acting on vectors in $\mathbb{R}^m$

$$\boldsymbol{\alpha}^* = (\alpha_2/\alpha_1, \ldots, \alpha_m/\alpha_1)^T, \qquad \boldsymbol{\alpha} \in \mathbb{R}^m,\ \alpha_1 > 0.$$

Note that $\boldsymbol{a}^* = (a_2, \ldots, a_m)^T$ because $a_1 = 1$ and thus $\langle f, \boldsymbol{a}\rangle = \boldsymbol{\rho}$ implies $\langle f^*, \boldsymbol{a}^*\rangle = \boldsymbol{\rho}^*$.
Apart from the passage to probability measures, we consider the functional of relative entropy. If P and R are probability measures on the Borel sets $\mathcal{B}$ on $\mathbb{R}^d$, such that P has a density with respect to R, i.e.

$$P(A) = \int_A p_R\, dR, \qquad A \in \mathcal{B}$$

the relative entropy (or I-divergence) is defined as

$$I(P||R) = \int p_R \log p_R \, dR.$$

As measure R we are going to use

$$R(A) = \int_A g^* \, d\boldsymbol{k}, \qquad g(\boldsymbol{k}) = \exp(-\mathcal{E}(\boldsymbol{k})) \tag{13}$$

where g is integrable since $\mathcal{E}(\boldsymbol{k})$ grows linearly (see (7)). Then, if P_{f^*} has density $f^* \in \mathcal{F}^*$ with respect to the Lebesgue measure, it has density f^*/g^* with respect to R and

$$I(P_{f^*}||R) = \int \frac{f^*}{g^*} \log \frac{f^*}{g^*} \, dR = \int f^* \log \frac{f^*}{g^*} \, d\boldsymbol{k}$$

Using the definition (11) of H and

$$\log f^* = \log f - \log \langle f, 1 \rangle , \qquad \log g^* = -\mathcal{E} - \log \langle g, 1 \rangle ,$$

we obtain the relation

$$I(P_{f^*}||R) = -\frac{1}{\langle f, 1 \rangle} H(f) + 1 + \log \frac{\langle g, 1 \rangle}{\langle f, 1 \rangle} + \frac{\langle f, \mathcal{E} \rangle}{\langle f, 1 \rangle}. \tag{14}$$

Since $\langle f, 1 \rangle$ and $\langle f, \mathcal{E} \rangle$ are constant on the set of densities $f \in \mathcal{F}$ with $\langle f, \boldsymbol{a} \rangle = \boldsymbol{\rho}$, we see that maximizing H subject to $\langle f, \boldsymbol{a} \rangle = \boldsymbol{\rho}$ is equivalent to

$$\begin{gathered} \text{minimize } I(P_{f^*}||R) \\ \text{subject to } f^* \in \mathcal{F}^* \text{ and } \langle f^*, \boldsymbol{a}^* \rangle = \boldsymbol{\rho}^* \end{gathered} \tag{15}$$

In summary, we have

Proposition 2 *Let $\boldsymbol{\rho} \in \mathcal{M}$. Then problem* (12) *has a unique solution $f \in \mathcal{F}$ if and only if* (15) *has a unique solution $f^* \in \mathcal{F}^*$. The relation between f and f^* is given by $f = \rho_1 f^*$. In particular, if $f^* = c\exp(\boldsymbol{\xi} \cdot \boldsymbol{a}^*)$ for some $\boldsymbol{\xi} \in \mathbb{R}^{m-1}$ and some $c > 0$, then $f = \exp(\boldsymbol{\lambda} \cdot \boldsymbol{a})$ with $\boldsymbol{\lambda} = (\log(c\rho_1), \xi_1, \ldots, \xi_{m-1})^T$.*

Csiszar's result, which is presented in the next section, shows that (15) is even uniquely solvable with an exponential density if P_{f^*} is replaced by general probability measures P on $\mathbb{R}^d$ which have the correct moments. In connection with Proposition 2 this immediately yields Theorem 1.

3.2 A general result by Csiszar

Csiszar's result [28] applies to general measurable spaces $(X, \mathcal{H})$ with weight functions $\boldsymbol{a}^* = (a_2, \ldots, a_m)$ being $\mathcal{H}$-measurable. Note that a_i^* can be general measurable functions here. We only keep the previous notation to be consistent with our use of Csiszar's theorem. By $\mathcal{P}$ we denote the set of probability measures on $(X, \mathcal{H})$ and for $P, R \in \mathcal{P}$, we write $P \ll R$ if P has an R-density, i.e.

$$P(A) = \int_A p_R \, dR, \qquad A \in \mathcal{H}$$

The entropy of P relative to R is defined as

$$I(P||R) = \begin{cases} \int p_R \log p_R \, dR & P \ll R \\ +\infty & P \not\ll R \end{cases}$$

To state Csiszar's theorem, we introduce for any $\boldsymbol{\rho}^* \in \mathbb{R}^{m-1}$ the set of all measures with moments $\boldsymbol{\rho}^*$

$$E(\boldsymbol{\rho}^*) = \{P \in \mathcal{P} \, : \, \boldsymbol{a}^* \in \mathbb{L}^1(P), \; \int \boldsymbol{a}^* \, dP = \boldsymbol{\rho}^*\}.$$

Since our aim is to find a minimizer of the relative entropy $I(P||R)$ over $P \in E(\boldsymbol{\rho}^*)$ for some fixed $R \in \mathcal{P}$, we restrict ourselves to those moment vectors $\boldsymbol{\rho}^*$ for which at least one corresponding $P \in \mathcal{P}$ has finite relative entropy, i.e.

$$A_R = \{\boldsymbol{\rho}^* \in \mathbb{R}^{m-1} \, : \, \exists P \in E(\boldsymbol{\rho}^*) \text{ such that } I(P||R) < \infty\}.$$

The last ingredient is the set T_R which contains all coefficient vectors $\boldsymbol{\xi} \in \mathbb{R}^{m-1}$ for which the density $\exp(\boldsymbol{\xi} \cdot \boldsymbol{a}^*)$ is R-integrable

$$T_R = \{\boldsymbol{\xi} \in \mathbb{R}^{m-1} \, : \, \exp(\boldsymbol{\xi} \cdot \boldsymbol{a}^*) \in \mathbb{L}^1(R)\}.$$

Using this notation, Theorem 3.3 of [28] can be formulated as

Theorem 3 *Assume T_R is open and let $\boldsymbol{\rho}^* \in \mathrm{int} A_R$. Then the problem*

$$\min_{P \in E(\boldsymbol{\rho}^*)} I(P||R)$$

has a unique solution $P \in \mathcal{P}$ with $P \ll R$ and density $p_R = c \exp(\boldsymbol{\xi} \cdot \boldsymbol{a}^*)$ for some vector $\boldsymbol{\xi} \in \mathbb{R}^{m-1}$ and some $c > 0$.

In order to apply Theorem 3 to our particular choice of weight functions and moment vectors, we just have to check its assumptions. In the following sections, we show that for any $\boldsymbol{\rho} \in \mathcal{M}$ there exists $f \in \mathcal{F}$ such that $\langle f, \boldsymbol{a} \rangle = \boldsymbol{\rho}$ and $f \log f \in \mathbb{L}^1(\mathbb{R}^d)$. In view of (14) this implies that the set of all normalized moments

$$\mathcal{M}^* = \{\langle f^*, \boldsymbol{a}^* \rangle \, : \, f^* \in \mathcal{F}^*\}$$

is contained in A_R and since $\mathcal{M}^*$ is open (Corollary 8), we have $\mathcal{M}^* \subset \mathrm{int} A_R$. The remaining condition that T_R is open is shown in Proposition 12. While the condition $\boldsymbol{\rho}^* \in \mathrm{int} A_R$ follows from very basic properties of the weight functions (linear independence and analyticity), the condition that T_R is open requires a detailed investigation of the growth behavior of $\boldsymbol{a}(\boldsymbol{k})$ for $|\boldsymbol{k}| \to \infty$.
Finally, we want to remark that Theorem 3 is not helpful in the parabolic band approximation where $\mathcal{E}(\boldsymbol{k}) = |\boldsymbol{k}|^2/2$ and the weights are simply polynomials in $\boldsymbol{k}$ where at least one of them grows faster than quadratic, say like $|\boldsymbol{k}|^4$. While it is still possible to show $\boldsymbol{\rho}^* \in \mathrm{int} A_R$, it is generally not true that T_R is open, essentially because $\exp(-|\boldsymbol{k}|^2/2 + 0 \cdot |\boldsymbol{k}|^4) \in \mathbb{L}^1$ but $\exp(-|\boldsymbol{k}|^2/2 + \epsilon |\boldsymbol{k}|^4) \notin \mathbb{L}^1$ for any $\epsilon > 0$ (see [6, 7] for details).

3.3 The weight functions

Before considering the moment set $\mathcal{M}$ in more detail, we collect a few basic properties of the weight functions. Our first result concerns the linear independence.

Proposition 4 *The components of the vector of weight functions* $\boldsymbol{a}$ *defined in* (8) *are linearly independent.*

Proof: Assuming the opposite, we could find $\gamma_i \in \mathbb{R}$ with $\sum |\gamma_i| > 0$ such that $\sum \gamma_i a_i(\boldsymbol{k}) = 0$ for all $\boldsymbol{k} \in \mathbb{R}^d$. Setting

$$P(\boldsymbol{v}) = \sum_{i=1}^{m_1} \gamma_i P_i(\boldsymbol{v}), \qquad Q(\boldsymbol{v}) = \sum_{i=1}^{m_2} \gamma_{i+m_1} Q_i(\boldsymbol{v})$$

this implies

$$P(\boldsymbol{v}(\boldsymbol{k})) + \mathcal{E}(\boldsymbol{k}) Q(\boldsymbol{v}(\boldsymbol{k})) = 0, \qquad \forall \boldsymbol{k} \in \mathbb{R}^d. \tag{16}$$

If we show that (16) implies $P = Q = 0$, the linear independence of P_i and Q_i leads to $\sum |\gamma_i| = 0$, in contradiction to the assumption. To show $P = Q = 0$, we take any $\boldsymbol{e} \in \mathbb{R}^d$ with $|\boldsymbol{e}| = 1$ and set

$$p(s) = P(s\boldsymbol{e}), \qquad q(s) = Q(s\boldsymbol{e}), \qquad s \in \mathbb{R}.$$

Note that $\boldsymbol{v}(s\boldsymbol{e}) = s(1+s^2)^{-1/2}\boldsymbol{e}$ and $\epsilon(s) = \mathcal{E}(s\boldsymbol{e}) = \sqrt{1+s^2} - 1$. Relation (16) implies

$$p\left(\frac{s}{\sqrt{1+s^2}}\right) + \epsilon(s) q\left(\frac{s}{\sqrt{1+s^2}}\right) = 0 \qquad \forall s \in \mathbb{R}.$$

Dividing by $\epsilon(s)$ and sending s to $\pm\infty$ yields $q(\pm 1) = 0$ so that $q(s) = (1-s^2)\tilde{q}(s)$. Hence

$$p\left(\frac{s}{\sqrt{1+s^2}}\right) + \epsilon(s) \frac{1}{1+s^2} \tilde{q}\left(\frac{s}{\sqrt{1+s^2}}\right) = 0, \qquad \forall s \in \mathbb{R}.$$

and by sending again s to $\pm\infty$, we find $p(\pm 1) = 0$, i.e. $p(s) = (1-s^2)\tilde{p}(s)$. Altogether, we get

$$\tilde{p}\left(\frac{s}{\sqrt{1+s^2}}\right) + \epsilon(s) \tilde{q}\left(\frac{s}{\sqrt{1+s^2}}\right) = 0, \qquad \forall s \in \mathbb{R}.$$

Repeating the argument, we conclude that $p = q = 0$ since otherwise the degree of p and q would be larger than any fixed number. Since we picked $\boldsymbol{e}$ arbitrarily, we also find $P = Q = 0$ which concludes the proof. ■

While linear independence only implies that the zero set $\{\boldsymbol{k} : \boldsymbol{\beta} \cdot \boldsymbol{a}(\boldsymbol{k}) = 0\}$ of a linear combination $\boldsymbol{\beta} \cdot \boldsymbol{a}$ of the weight functions *cannot be very big* (the whole space $\mathbb{R}^d$), we will need the stronger property that the zero set *must be very small* in the following sense.

Definition 5 *A set of measurable functions* $a_1, \ldots, a_m$ *on* $\mathbb{R}^d$ *has the pseudo-Haar property if for any* $0 \neq \boldsymbol{\beta} \in \mathbb{R}^m$*, the zero set of* $\boldsymbol{\beta} \cdot \boldsymbol{a}$ *has zero Lebesgue measure.*

Proposition 6 *The components of the vector of weight functions* $\boldsymbol{a}$ *defined in* (8) *have the pseudo-Haar property.*

Proof: We follow the argument presented in [29]: since $\mathcal{E}, \boldsymbol{v}, P_i, Q_i$ are analytic, also the weights a_i are analytic and since the zero set of any non-zero analytic function on $\mathbb{R}^d$ has vanishing Lebesgue measure, the linear independence of the weight functions implies the pseudo-Haar property. ■

3.4 The moment cone

Since $\mathcal{M} = \{\langle f, \boldsymbol{a}\rangle : f \in \mathcal{F}\}$ is the image of $\mathcal{F}$ under the linear mapping $f \mapsto \langle f, \boldsymbol{a}\rangle$, the obvious property of $\mathcal{F}$ being a convex cone carries over to $\mathcal{M}$.

Proposition 7 *The moment set $\mathcal{M}$ is an open convex cone in $\mathbb{R}^m$.*

Proof: It remains to show that $\mathcal{M}$ is open which we do by using the same argument as in [29]. Assuming that $\bar{\boldsymbol{\rho}} = \langle f, \boldsymbol{a}\rangle$ is a boundary point of $\mathcal{M}$, there exists $0 \neq \boldsymbol{\beta} \in \mathbb{R}^m$ such that $(\boldsymbol{\rho} - \bar{\boldsymbol{\rho}}) \cdot \boldsymbol{\beta} \geq 0$ for all $\boldsymbol{\rho} \in \mathcal{M}$ due to convexity. However, the function $h \in \mathcal{F}$ defined by

$$h(\boldsymbol{k}) = \begin{cases} \frac{3}{2} f(\boldsymbol{k}) & \boldsymbol{\beta} \cdot \boldsymbol{a}(\boldsymbol{k}) < 0 \\ \frac{1}{2} f(\boldsymbol{k}) & \boldsymbol{\beta} \cdot \boldsymbol{a}(\boldsymbol{k}) \geq 0 \end{cases}$$

has a moment vector $\boldsymbol{\rho} = \langle h, \boldsymbol{a}\rangle$ which satisfies

$$(\boldsymbol{\rho} - \bar{\boldsymbol{\rho}}) \cdot \boldsymbol{\beta} = \langle h - f, \boldsymbol{\beta} \cdot \boldsymbol{a}\rangle = -\frac{1}{2} \langle f, |\boldsymbol{\beta} \cdot \boldsymbol{a}|\rangle < 0$$

which is strictly negative because $|\boldsymbol{\beta} \cdot \boldsymbol{a}|$ can vanish at most on a set of measure zero due to the pseudo-Haar property. Hence, the assumption that $\bar{\boldsymbol{\rho}}$ is a boundary point leads to a contradiction and $\mathcal{M}$ is therefore open. ■

We remark that the set $\mathcal{M}^*$ of normalized moments is obtained by intersecting the cone $\mathcal{M}$ and the hyperplane $\{1\} \times \mathbb{R}^{m-1}$, i.e.

$$\mathcal{M} \cap \{1\} \times \mathbb{R}^{m-1} = \{1\} \times \mathcal{M}^*.$$

Consequently, $\mathcal{M}^*$ is also convex because both $\mathcal{M}$ and $\{1\} \times \mathbb{R}^{m-1}$ are convex. Since $\mathcal{M}$ is open in $\mathbb{R}^m$, the intersection is open in the relative topology of $\{1\} \times \mathbb{R}^{m-1}$ which is equivalent to the usual topology of $\mathbb{R}^{m-1}$ (up to the bijection $(1, \boldsymbol{\alpha}) \mapsto \boldsymbol{\alpha}$ between $\{1\} \times \mathbb{R}^{m-1}$ and $\mathbb{R}^{m-1}$).

Corollary 8 *The set of normalized moments $\mathcal{M}^*$ is open and convex in $\mathbb{R}^{m-1}$.*

3.5 The entropy functional

A minimal requirement for the maximum entropy problem (12) to have a solution is the existence of at least one density $f \in \mathcal{F}$ with $\langle f, \boldsymbol{a}\rangle = \boldsymbol{\rho}$ which has a *finite* entropy. This question is considered here.

Proposition 9 *Let $\boldsymbol{\rho} \in \mathcal{M}$. Then there exists a function $f \in \mathcal{F}$ with $\langle f, \boldsymbol{a}\rangle = \boldsymbol{\rho}$ and $f \log f \in \mathbb{L}^1(\mathbb{R}^d)$. In particular, f^* satisfies $\langle f^*, \boldsymbol{a}^*\rangle = \boldsymbol{\rho}^*$ and $f^* \log f^* \in \mathbb{L}^1(\mathbb{R}^d)$.*

Proof: Since $\mathcal{M}$ is open, there exists a hypercube with vertices $\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_{2^m}$ in $\mathcal{M}$ which has $\boldsymbol{\rho}$ as its center point. By definition of $\mathcal{M}$, each $\boldsymbol{\eta}_i$ is the moment vector of some $f_i \in \mathcal{F}$. For $j \in \mathbb{N}$, we set

$$f_i^{(j)}(\boldsymbol{k}) = \begin{cases} \min\{j, f_i(\boldsymbol{k})\} & |\boldsymbol{k}| < j \\ 0 & |\boldsymbol{k}| \geq j \end{cases}$$

Using the dominated convergence theorem, it then follows that

$$\boldsymbol{\eta}_i^{(j)} = \left\langle f_i^{(j)}, \boldsymbol{a} \right\rangle \xrightarrow[j \to \infty]{} \langle f_i, \boldsymbol{a} \rangle = \boldsymbol{\eta}_i.$$

Hence, for j large enough, the vectors $\boldsymbol{\eta}_i^{(j)}$ will also be in $\mathcal{M}$ and $\boldsymbol{\rho}$ will be in their convex hull., i.e. there exist $\omega_i \geq 0$ which add up to one, such that

$$\boldsymbol{\rho} = \sum_{i=1}^{2^m} \omega_i \boldsymbol{\eta}_i^{(j)}$$

(the argument is based on a simple application of the implicit function theorem – see [8] for details). Setting $f = \sum_i \omega_i f_i^{(j)}$, we thus have $\boldsymbol{\rho} = \langle f, \boldsymbol{a} \rangle$ with f being bounded and of compact support. Consequently, $f \log f \in \mathbb{L}^1(\mathbb{R}^d)$. The result for f^* follows by normalization. ■

3.6 The Lagrange multipliers

In [7, 8] it is shown that the topology of the set

$$\Lambda = \{\boldsymbol{\lambda} \in \mathbb{R}^m : \exp(\boldsymbol{\lambda} \cdot \boldsymbol{a}) \in \mathbb{L}^1(\mathbb{R}^d)\} \tag{17}$$

(the so called Lagrange multipliers) determines the solvability of the maximum entropy problem. Also in Csiszar's theorem, it is decisive that the set

$$T_R = \{\boldsymbol{\xi} \in \mathbb{R}^{m-1} : \exp(\boldsymbol{\xi} \cdot \boldsymbol{a}^*) \in \mathbb{L}^1(R)\} \tag{18}$$

is open in $\mathbb{R}^{m-1}$. Note that Λ and T_R are closely related if we choose the probability measure R according to (13). In fact, if $\boldsymbol{\xi} \in T_R$ then $\exp(\boldsymbol{\xi} \cdot \boldsymbol{a}^*) \exp(-\mathcal{E}) \in \mathbb{L}^1(\mathbb{R}^d)$ so that the vector $H(\boldsymbol{\xi})$ defined by $H(\boldsymbol{\xi}) \cdot \boldsymbol{a} = \boldsymbol{\xi} \cdot \boldsymbol{a}^* - \mathcal{E}$ is contained in Λ. Note that $H(\boldsymbol{\xi}) = B\boldsymbol{\xi} - \boldsymbol{b}$ is an affine linear mapping where the vector $\boldsymbol{b}$ is the unit vector which picks out the component $\mathcal{E}$ of the weight vector, i.e. $\boldsymbol{b} \cdot \boldsymbol{a} = \mathcal{E}$ and B is the canonical embedding operator of $\mathbb{R}^{m-1}$ into $\mathbb{R}^m$, i.e. $B\boldsymbol{\xi} = (0, \xi_1, \dots, \xi_{m-1})^T$. Conversely, $H(\boldsymbol{\xi}) \in \Lambda$ implies that $\boldsymbol{\xi} \in T_R$ (by multiplying $H(\boldsymbol{\xi}) = \boldsymbol{\lambda}$ by B^T and using that $B^T B$ is the identity on $\mathbb{R}^{m-1}$). Altogether, we conclude that T_R is the pre-image of Λ and since H is continuous we find that T_R is open in $\mathbb{R}^{m-1}$ if Λ is open in $\mathbb{R}^m$. In the following, we therefore restrict our considerations to a characterization of Λ.

Introducing the polynomial vectors $\boldsymbol{P} = (P_1, \dots, P_{m_1})^T$, $\boldsymbol{Q} = (Q_1, \dots, Q_{m_2})^T$ and splitting $\boldsymbol{\lambda} \in \mathbb{R}^m$ into $\boldsymbol{\lambda}_1 \in \mathbb{R}^{m_1}$, $\boldsymbol{\lambda}_2 \in \mathbb{R}^{m_2}$, we have by definition of the weight functions

$$\boldsymbol{\lambda} \cdot \boldsymbol{a} = \boldsymbol{\lambda}_1 \cdot \boldsymbol{P}(\boldsymbol{v}) + \boldsymbol{\lambda}_2 \cdot \boldsymbol{Q}(\boldsymbol{v})\mathcal{E}.$$

In view of (7), the velocity $\boldsymbol{v}(\boldsymbol{k})$ is a bounded function of $\boldsymbol{k}$ and thus also $\boldsymbol{P}(\boldsymbol{v}(\boldsymbol{k}))$ and $\boldsymbol{Q}(\boldsymbol{v}(\boldsymbol{k}))$. Integrability of $\exp(\boldsymbol{\lambda}\cdot\boldsymbol{a})$ can therefore only be achieved if the factor $\boldsymbol{\lambda}_2\cdot\boldsymbol{Q}(\boldsymbol{v})$ in front of $\mathcal{E}$ is uniformly negative for large $\boldsymbol{k}$ since $\mathcal{E}$ grows linearly (see (7)). Observing that $\boldsymbol{v}(\boldsymbol{k})$ tends to the unit sphere for $|\boldsymbol{k}|\to\infty$, this leads to the integrability condition on $\boldsymbol{\lambda}_2$

$$\boldsymbol{\lambda}_2\cdot\boldsymbol{Q}(\boldsymbol{e})<0 \qquad \forall \boldsymbol{e}\in S_{d-1}, \tag{19}$$

where S_{d-1} is the unit sphere of $\mathbb{R}^d$. With the following result, we make our considerations more precise (for ease of notation, we identify $(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2)\in\mathbb{R}^{m_1}\times\mathbb{R}^{m_2}$ with $\binom{\boldsymbol{\lambda}_1}{\boldsymbol{\lambda}_2}\in\mathbb{R}^{m_1+m_2}$).

Lemma 10 *Let $\boldsymbol{\lambda}_1\in\mathbb{R}^{m_1}$ and $\boldsymbol{\lambda}_2\in\mathbb{R}^{m_2}$ such that* (19) *is satisfied. Then $(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2)\in\Lambda$.*

Proof: Assume $\boldsymbol{\lambda}_2\cdot\boldsymbol{Q}(\boldsymbol{e})<0$ for all $|\boldsymbol{e}|=1$. Then there exists $\mu>0$ such that $\boldsymbol{\lambda}_2\cdot\boldsymbol{Q}(\boldsymbol{e})\le-\mu$ for all $|\boldsymbol{e}|=1$. In particular, we can find $\delta>0$ such that

$$\boldsymbol{\lambda}_2\cdot\boldsymbol{Q}(\boldsymbol{v})\le-\mu/2, \qquad 1-\delta\le|\boldsymbol{v}|\le 1$$

and we remark that, in view of (6),

$$|\boldsymbol{v}|\ge 1-\delta \qquad \Leftrightarrow \qquad |\boldsymbol{k}|\ge\frac{1-\delta}{\sqrt{1-(1-\delta)^2}}=D.$$

Since $|\boldsymbol{v}(\boldsymbol{k})|<1$ for all $\boldsymbol{k}\in\mathbb{R}^d$, we have

$$|\boldsymbol{\lambda}_1\cdot\boldsymbol{a}_1(\boldsymbol{k})|\le|\boldsymbol{\lambda}_1|\max_{|\boldsymbol{v}|\le 1}|\boldsymbol{P}(\boldsymbol{v})|=C_1$$

and thus

$$\boldsymbol{\lambda}\cdot\boldsymbol{a}(\boldsymbol{k})\le C_1-\frac{\mu}{2}\mathcal{E}(\boldsymbol{k}), \qquad |\boldsymbol{k}|\ge D.$$

On $|\boldsymbol{k}|\le D$ we can find uniform bounds so that

$$\boldsymbol{\lambda}\cdot\boldsymbol{a}(\boldsymbol{k})\le C_2-\frac{\mu}{2}\mathcal{E}(\boldsymbol{k}), \qquad \boldsymbol{k}\in\mathbb{R}^d$$

with a suitable $C_2>0$. Taking into account that for $|\boldsymbol{k}|\ge 2$

$$\mathcal{E}(\boldsymbol{k})=|\boldsymbol{k}|\sqrt{1/|\boldsymbol{k}|^2+1}-1\ge|\boldsymbol{k}|-|\boldsymbol{k}|/2=\frac{|\boldsymbol{k}|}{2}$$

we eventually get the estimate

$$\boldsymbol{\lambda}\cdot\boldsymbol{a}(\boldsymbol{k})\le C-\nu|\boldsymbol{k}|, \qquad \boldsymbol{k}\in\mathbb{R}^d$$

for some $C,\nu>0$ which yields $\exp(\boldsymbol{\lambda}\cdot\boldsymbol{a}(\boldsymbol{k}))\le\exp(C-\nu|\boldsymbol{k}|)$, i.e. $\boldsymbol{\lambda}\in\Lambda$. ∎

The next result shows that (19) is also necessary for integrability.

Lemma 11 *Let $\boldsymbol{\lambda}_1\in\mathbb{R}^{m_1}$ and $\boldsymbol{\lambda}_2\in\mathbb{R}^{m_2}$ such that* (19) *is violated. Then $(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2)\notin\Lambda$.*

Proof: Our aim is to show that $\boldsymbol{\lambda} \cdot \boldsymbol{a}$ is bounded from below on a set $\Omega_{\bar{s}}$ of infinite measure. If $Q(\boldsymbol{v}) = \boldsymbol{\lambda}_2 \cdot \boldsymbol{Q}(\boldsymbol{v})$ satisfies $Q(\boldsymbol{e}_0) \geq 0$ with $|\boldsymbol{e}_0| = 1$, we set

$$\Omega_{\bar{s}} = \{s\boldsymbol{e}_0 + \boldsymbol{\delta} : s \geq \bar{s},\ |\boldsymbol{\delta}| \leq 1,\ \boldsymbol{\delta} \cdot \boldsymbol{e}_0 = 0\}$$

which is a cylinder of radius one around the direction $\boldsymbol{e}_0$. In particular, $|\Omega_{\bar{s}}| = \infty$ for all $\bar{s} \geq 0$. An obvious parameterization of $\Omega_{\bar{s}}$ is given by $[\bar{s}, \infty) \times B$ where $B = \{\boldsymbol{\delta} \in \mathbb{R}^d : |\boldsymbol{\delta}| \leq 1,\ \boldsymbol{\delta} \cdot \boldsymbol{e}_0 = 0\}$. Then $\boldsymbol{k}(s, \boldsymbol{\delta}) = s\boldsymbol{e}_0 + \boldsymbol{\delta}$ and $s(\boldsymbol{k}) = \boldsymbol{k} \cdot \boldsymbol{e}_0$, $\boldsymbol{\delta}(\boldsymbol{k}) = \boldsymbol{k} - (\boldsymbol{k} \cdot \boldsymbol{e}_0)\boldsymbol{e}_0$. For $|\boldsymbol{k}|$ we have the estimate

$$s^2 \leq |\boldsymbol{k}(s, \boldsymbol{\delta})|^2 \leq s^2 + 1 \tag{20}$$

respectively, for $\bar{s} \geq 1$,

$$s \leq |\boldsymbol{k}(s, \boldsymbol{\delta})| \leq 2s. \tag{21}$$

Using the estimates $|\boldsymbol{k}|/2 \leq \mathcal{E}(\boldsymbol{k}) \leq 2(1 + |\boldsymbol{k}|)$ for $|\boldsymbol{k}| \geq 2$ from the proof of Lemma 10, we conclude that

$$s/2 \leq \mathcal{E}(\boldsymbol{k}(s, \boldsymbol{\delta})) \leq 4(1 + s) \qquad s \geq 2, \boldsymbol{\delta} \in B. \tag{22}$$

Using (20) and the definition of $\boldsymbol{v}$, we find

$$0 \leq 1 - |\boldsymbol{v}(\boldsymbol{k}(s, \boldsymbol{\delta}))|^2 \leq 1 - \frac{s^2}{s^2 + 2} = 1 - \frac{1}{1 + \frac{2}{s^2}}.$$

An elementary estimate shows $1 - 1/(1 + x) \leq x$ for $-1 < x \leq 1$ and hence

$$0 \leq 1 - |\boldsymbol{v}(\boldsymbol{k}(s, \boldsymbol{\delta}))|^2 \leq \frac{2}{s^2} \qquad s \geq 2, \boldsymbol{\delta} \in B. \tag{23}$$

Similarly, we find

$$0 \leq 1 - \boldsymbol{v}(\boldsymbol{k}(s, \boldsymbol{\delta})) \cdot \boldsymbol{e}_0 \leq 1 - \frac{s}{\sqrt{2 + s^2}}$$

and the elementary estimate $1 - 1/\sqrt{1 + x} \leq x$ for $0 \leq x \leq 1$ yields

$$0 \leq 1 - \boldsymbol{v}(\boldsymbol{k}(s, \boldsymbol{\delta})) \cdot \boldsymbol{e}_0 \leq \frac{2}{s^2} \qquad s \geq 2, \boldsymbol{\delta} \in B. \tag{24}$$

Using the relation $|\boldsymbol{v} - \boldsymbol{e}_0|^2 = |\boldsymbol{v}|^2 - 1 + 2(1 - \boldsymbol{v} \cdot \boldsymbol{e}_0)$ together with (24) and (23), we get

$$|\boldsymbol{v}(\boldsymbol{k}(s, \boldsymbol{\delta})) - \boldsymbol{e}_0| \leq \frac{\sqrt{8}}{s} \qquad s \geq 2, \boldsymbol{\delta} \in B. \tag{25}$$

In particular, for any $\epsilon > 0$, we can find $\bar{s}(\epsilon) > 2$ such that $|\boldsymbol{v}(\boldsymbol{k}) - \boldsymbol{e}_0| < \epsilon$ for all $\boldsymbol{k} \in \Omega_{\bar{s}(\epsilon)}$. Assuming first that $Q(\boldsymbol{e}_0) > 0$, we can find $\epsilon > 0$ such that $Q(\boldsymbol{v}) \geq 0$ for all $|\boldsymbol{v} - \boldsymbol{e}_0| < \epsilon$. Hence, $Q(\boldsymbol{v}(\boldsymbol{k}))\mathcal{E}(\boldsymbol{k}) \geq 0$ for all $\boldsymbol{k}$ in the cylinder $\Omega_{\bar{s}(\epsilon)}$ and with $C_1 = \max_{|\boldsymbol{v}| \leq 1} |\boldsymbol{\lambda}_1 \cdot \boldsymbol{P}(\boldsymbol{v})|$, we conclude

$$\exp(\boldsymbol{\lambda} \cdot \boldsymbol{a}) \geq \exp(-C_1), \qquad \forall \boldsymbol{k} \in \Omega_{\bar{s}(\epsilon)}$$

which implies $\boldsymbol{\lambda} \notin \Lambda$ since $\Omega_{\bar{s}(\epsilon)}$ has infinite measure.
In the case that $\boldsymbol{e}_0$ is a root of Q, we can find $\epsilon > 0$ and $C > 0$ such that

$$|Q(\boldsymbol{v})| \leq C|\boldsymbol{v} - \boldsymbol{e}_0| \qquad \text{if } |\boldsymbol{v} - \boldsymbol{e}_0| < \epsilon. \tag{26}$$

Hence, on $\Omega_{\bar{s}(\epsilon)}$, we have with (22), (25), (26) $C_1 = \max_{|\boldsymbol{v}|\le 1} |\boldsymbol{\lambda}_1 \cdot \boldsymbol{P}(\boldsymbol{v})|$

$$\boldsymbol{\lambda} \cdot \boldsymbol{a}(\boldsymbol{k}(s,\boldsymbol{\delta})) \ge -C_1 - 4(1+s)C\frac{\sqrt{8}}{s} \ge -C_1 - 8\sqrt{8}C = -K \qquad s \ge \bar{s}(\epsilon) \ge 2.$$

Again $\exp(\boldsymbol{\lambda} \cdot \boldsymbol{a}) \ge \exp(-K)$ on the set $\Omega_{\bar{s}(\epsilon)}$ of infinite measure implies $\boldsymbol{\lambda} \notin \Lambda$. ■

Altogether, we have shown that Λ coincides with the set

$$C = \{(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) : \boldsymbol{\lambda}_1 \in \mathbb{R}^{m_1}, \boldsymbol{\lambda}_2 \in \mathbb{R}^{m_2},\ \boldsymbol{\lambda}_2 \cdot \boldsymbol{Q}(\boldsymbol{e}) < 0\ \forall |\boldsymbol{e}| = 1\}. \tag{27}$$

Proposition 12 *The set Λ defined in* (17) *is a non-empty, open, convex cone in $\mathbb{R}^m$. In particular, T_R defined in* (18) *with R given by* (13) *is open in $\mathbb{R}^{m-1}$.*

Proof: In view of our considerations above, it suffices to show that C defined in (27) is a non-empty, open, convex cone. Since $Q_1 = 1$, we find $(\boldsymbol{0}, -\boldsymbol{e}_1^{(m_2)}) \in C \neq \emptyset$, where $\boldsymbol{e}_1^{(m_2)}$ is the first unit vector in $\mathbb{R}^{m_2}$. Also convexity follows easily from the definition. Finally, if $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) \in C$, we can find $\mu > 0$ such that $\max_{|\boldsymbol{e}|=1} \boldsymbol{\lambda}_2 \cdot \boldsymbol{Q}(\boldsymbol{e}) \le -\mu$. Since $\boldsymbol{Q}(\boldsymbol{e})$ is bounded for $|\boldsymbol{e}| = 1$, we also find $\delta > 0$ such that $|\boldsymbol{\delta}_2 \cdot \boldsymbol{Q}(\boldsymbol{e})| < \mu/2$ for all $\boldsymbol{\delta}_2 \in \mathbb{R}^{m_2}$ with $|\boldsymbol{\delta}_2| < \delta$. Hence $(\boldsymbol{\lambda}_1 + \boldsymbol{\delta}_1, \boldsymbol{\lambda}_2 + \boldsymbol{\delta}_2) \in C$ for all $|\boldsymbol{\delta}_1|, |\boldsymbol{\delta}_2| < \delta$ which shows that C is open. ■

3.7 Proof of Theorem 1

In view of Proposition 2 it suffices to show that, for given $\boldsymbol{\rho} \in \mathcal{M}$, problem (15) has a unique solution of exponential type. Since $\boldsymbol{\rho} \in \mathcal{M}$ implies $\boldsymbol{\rho}^* \in \mathcal{M}^*$, Corollary 8 and Proposition 9 in connection with (14) can be used to show that $\boldsymbol{\rho}^* \in \mathrm{int}A_R$. Finally, T_R is open according to Proposition 12 and Theorem 3 shows that (15) has a unique solution with R-density $c\exp(\boldsymbol{\xi} \cdot \boldsymbol{a}^*)$, or equivalently, with Lebesgue density

$$c\exp(\boldsymbol{\xi} \cdot \boldsymbol{a}^*)\frac{\exp(-\mathcal{E})}{\langle \exp(-\mathcal{E}), 1\rangle} = \exp(\boldsymbol{\lambda} \cdot \boldsymbol{a})$$

where

$$\boldsymbol{\lambda} = (\log(c/\langle \exp(-\mathcal{E}), 1\rangle), \xi_1, \ldots, \xi_{m-1})^T.$$

This concludes the proof of Theorem 1. ■

REMARK. From the proof of Theorem 1 one can see that the boundedness of $|\boldsymbol{v}(\boldsymbol{k})|$ is the important property. Therefore a similar result can be expected for more general dispersion relations which exhibit an effect of saturation for the modulus of $\boldsymbol{v}(\boldsymbol{k})$.

4 The moment cone for the 5-moment and 8-moment MEP system

In this section we analyze in detail some specific MEP systems of relevance in the applications. The main aim is to give a numerical indication about the moment cone which represents the physical region where the solutions of the MEP systems can be found.

First the 5-moment system representing the Euler-Poisson model is studied, eventually the hydrodynamical one based on the macroscopic electron density, velocity, energy and energy-flux is analyzed. In order to make the results physically more transparent the original variables instead of the scaled one are again introduced. The numerical results are concerned with silicon and therefore $\alpha = 0.5$ while for the effective mass we use the value $m^* = 0.32 m_e$ with m_e being the electron mass in vacuum.

4.1 The Euler-Poisson model

It is based on the same moments employed in ideal gas dynamics, that is density n, average velocity $\boldsymbol{u}$ and average energy W. The resulting balance equations are

$$\frac{\partial n}{\partial t} + \frac{\partial (n u^i)}{\partial x^i} = 0, \tag{28}$$

$$\frac{\partial (n u^i)}{\partial t} + \frac{\partial (n U^{ij})}{\partial x^j} = -enE_j H^{ij} + nC_u^i, \tag{29}$$

$$\frac{\partial (nW)}{\partial t} + \frac{\partial (n S^j)}{\partial x^j} = -neu_k E^k + nC_W, \tag{30}$$

where

$$\begin{aligned} U^{ij} &= \frac{1}{n}\int_{\mathbb{R}^3} f v^i v^j d\boldsymbol{k}, \quad H^{ij} = \frac{1}{n}\int_{\mathbb{R}^3} \frac{1}{\hbar} f \frac{\partial v_i}{\partial k_j} d\boldsymbol{k}, \\ C_u^i &= \frac{1}{n}\int_{\mathbb{R}^3} \mathcal{C}[f] v^i d\boldsymbol{k}, \quad C_W = \frac{1}{n}\int_{\mathbb{R}^3} \mathcal{C}[f]\mathcal{E}(k) d\mathbf{k}. \end{aligned}$$

Of course in the parabolic band approximation the left hand side of the previous equations is the same of that arising for monatomic gases.
For the 5-moment case the weight function vector is

$$\boldsymbol{a} = (1, \boldsymbol{v}, \mathcal{E})$$

and the corresponding Lagrange multipliers are given by the vector

$$\boldsymbol{\lambda} = -\left(\lambda, \boldsymbol{\lambda}^v, \lambda^W\right).$$

The MEP distribution function reads

$$f_{\boldsymbol{\lambda}} = \exp\left(-\lambda - \lambda_i^v v^i - \lambda^W \mathcal{E}\right) \tag{31}$$

and one has the straightforward characterization of the cone Λ

$$\Lambda = \left\{\boldsymbol{\lambda} = -\left(\lambda, \boldsymbol{\lambda}^v, \lambda^W\right) : \ \boldsymbol{\lambda} \in \mathbb{R}^5, \lambda^W > 0\right\}$$

which is obviously convex and open. We remark that

$$\frac{\partial v_i}{\partial k_j} = \frac{\hbar}{m^*}\left[\frac{\delta_{ij}}{\sqrt{1 + \frac{2\alpha}{m^*}\hbar^2 |\boldsymbol{k}|^2}} - \frac{\frac{2\alpha}{m^*}\hbar^2 k_i k_j}{\left[1 + \frac{2\alpha}{m^*}\hbar^2 |\boldsymbol{k}|^2\right]^{3/2}}\right]$$

are bounded regular functions so that the moments H^{ij} are well defined for each $f_{\boldsymbol{\lambda}}$ with $\boldsymbol{\lambda} \in \Lambda$.

By choosing a reference frame $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ adapted to $\boldsymbol{u}$, one can write $\boldsymbol{u} = u_3\mathbf{e}_3$ and $\boldsymbol{\lambda}^v = \lambda_3^v\mathbf{e}_3$ while in spherical coordinates $\mathbf{v}(\boldsymbol{k}) = v(\mathcal{E})(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, where

$$v(\mathcal{E}) = \sqrt{\frac{2\mathcal{E}(1+\alpha\mathcal{E})}{m^*(1+2\alpha\mathcal{E})^2}} \tag{32}$$

is the modulus of $\mathbf{v}$ in terms of $\mathcal{E}$.
By writing

$$d\boldsymbol{k} = \frac{m^*}{\hbar^3}\sqrt{2m^*\mathcal{E}(1+\alpha\mathcal{E})}(1+2\alpha\mathcal{E})d\mathcal{E}\, d\Omega$$

with elementary solid angle $d\Omega$, the explicit relations between the Lagrange multipliers and the macroscopic variables are given by

$$\begin{aligned}
n &= 4\sqrt{2}\pi\frac{(m^*)^{3/2}}{\hbar^3}\, e^{-\lambda} d_0 && (33)\\
u_3 &= \frac{1}{d_0}\int_0^\infty v(\mathcal{E})e^{-\lambda^W\mathcal{E}}\sqrt{\mathcal{E}(1+\alpha\mathcal{E})}(1+2\alpha\mathcal{E})\left[\frac{\sinh z}{z^2} - \frac{\cosh z}{z}\right] d\mathcal{E} && (34)\\
W &= \frac{1}{d_0}\int_0^\infty \mathcal{E}e^{-\lambda^W\mathcal{E}}\sqrt{\mathcal{E}(1+\alpha\mathcal{E})}(1+2\alpha\mathcal{E})\ \frac{\sinh z}{z}\, d\mathcal{E} && (35)
\end{aligned}$$

with

$$\begin{aligned}
z &= \lambda_3^v\, v(\mathcal{E}), && (36)\\
d_0 &= \int_0^\infty e^{-\lambda^W\mathcal{E}}\sqrt{\mathcal{E}(1+\alpha\mathcal{E})}(1+2\alpha\mathcal{E})\ \frac{\sinh z}{z}\, d\mathcal{E} && (37)
\end{aligned}$$

Practically there is not limitation on n. Therefore it is relevant only to study the dependence of u_3 and W on λ_3^v and λ^W. In order to have an estimate about the moment cone $\mathcal{M}$ we have numerically evaluated the relations (34)-(35). We remark that as $z \mapsto 0$ some numerical difficulty arises which can be easily overcome by using for $|z| \ll 1$ the Taylor expansions

$$\begin{aligned}
\frac{\sinh z}{z} &= 1 + \frac{1}{6}z^2 + \frac{1}{120}z^4 + O(z^5),\\
\frac{\sinh z}{z^2} - \frac{\cosh z}{z} &= -\frac{1}{3}z - \frac{1}{30}z^3 + O(z^5).
\end{aligned}$$

In figure 1 there is plotted the image of the rectangle $\left\{(\lambda_3^v/\sqrt{m^*}, \lambda^W) \in [-10, 10] \times [1, 65]\right\}$ under the mapping $(\lambda_3^v, \lambda^W) \mapsto (u_3, W)$ defined by the relations (34)-(35). $\lambda_3^v/\sqrt{m^*}$ is expressed in $1/\sqrt{eV}$ and λ^W in $1/eV$.
In electron device simulations typical values of the energy are between 0.03 eV and 0.5 eV and typical values of the average velocity u are between zero and few times the saturation velocity v_S which is about 10^5 m/sec in silicon. The figure 1 shows that the moment cone is sufficiently wide to enclose the relevant physical region of density, velocity and energy.

4.2 The 8-moment model

The 8-moment model describes the electron as a heat-conducing gas and has been considered in several articles [1, 14, 21, 22] where the relations between Lagrange multipliers

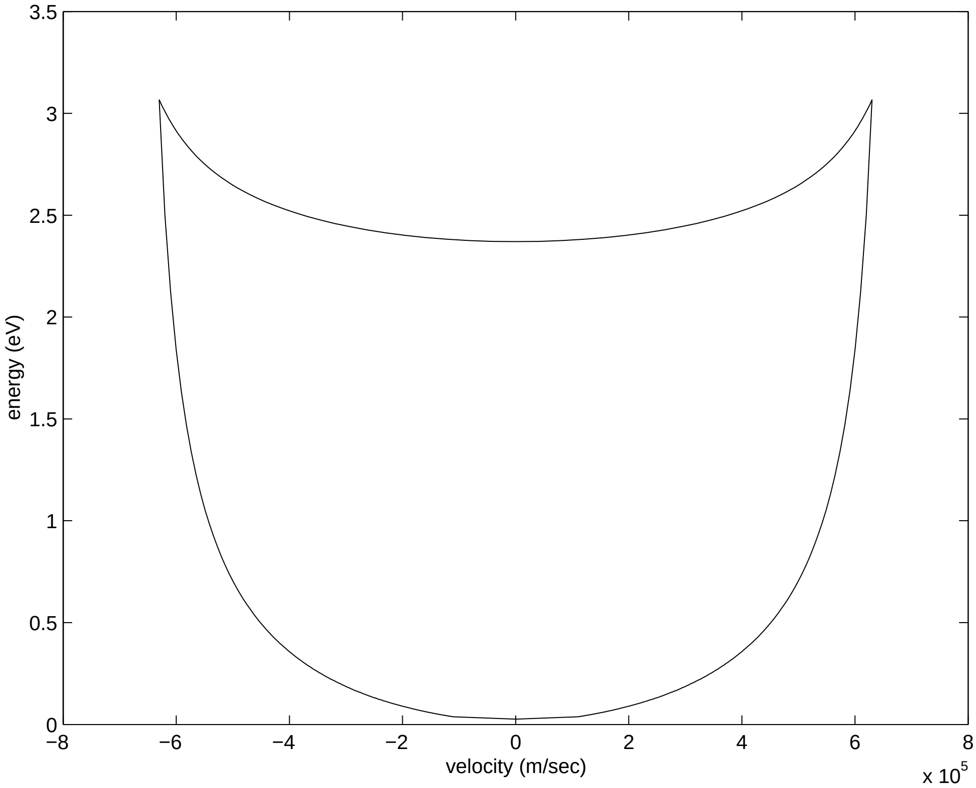

Figure 1: image of the rectangle $\left\{(\lambda_3^v/\sqrt{m^*}, \lambda^W) \in [-10, 10] \times [1, 65]\right\}$ under the mapping $(\lambda_3^v, \lambda^W) \mapsto (u_3, W)$ defined by the relations (34)-(35).

and basic moments are obtained by expansions in terms of a small anisotropy parameter. Here the analysis of the dependence on the multipliers is studied for the exact closure.
As fundamental macroscopic variables one takes the density n, the average velocity $\boldsymbol{u}$, the average energy W and the average energy-flux $\boldsymbol{S}$. The resulting balance equations are

$$\frac{\partial n}{\partial t} + \frac{\partial (n u^i)}{\partial x^i} = 0, \tag{38}$$

$$\frac{\partial (n u^i)}{\partial t} + \frac{\partial (n U^{ij})}{\partial x^j} = -enE_j H^{ij} + nC_u^i, \tag{39}$$

$$\frac{\partial (nW)}{\partial t} + \frac{\partial (n S^j)}{\partial x^j} = -neu_k E^k + nC_W, \tag{40}$$

$$\frac{\partial (n S^i)}{\partial t} + \frac{\partial (n F^{ij})}{\partial x^j} + neE_j G^{ij} = nC_{S^i}, \tag{41}$$

where, with respect to the 5-moment case, the additional tensorial quantities are defined as

$$\begin{aligned} G^{ij} &= \frac{1}{n}\int_{\mathbb{R}^3} \frac{1}{\hbar} f \frac{\partial}{\partial k_j}(\mathcal{E} v_i) d\mathbf{k}, \\ F^{ij} &= \frac{1}{n}\int_{\mathbb{R}^3} f v^i v^j \mathcal{E}(k) d\mathbf{k}, \\ C_{S^i} &= \frac{1}{n}\int_{\mathbb{R}^3} \mathcal{C}[f] v^i \mathcal{E}(k) d\mathbf{k}. \end{aligned}$$

For the 8-moment case under consideration the weight function vector is

$$\boldsymbol{a} = (1, \boldsymbol{v}, \mathcal{E}, \mathcal{E}\boldsymbol{v})$$

and the corresponding Lagrange multipliers are given by the vector

$$\boldsymbol{\lambda} = -\left(\lambda, \boldsymbol{\lambda}^v, \lambda^W, \boldsymbol{\lambda}^W\right).$$

The MEP distribution function reads

$$f_{\boldsymbol{\lambda}} = \exp\left(-\lambda - \lambda_i^v v^i - \lambda^W \mathcal{E} - \lambda_i^W \mathcal{E} v^i\right) \tag{42}$$

and, by taking into account the dependence of the modulus of $\boldsymbol{v}$ on $\mathcal{E}$ (see fig. 2), one has the following characterization of the moment cone Λ

$$\Lambda = \left\{\boldsymbol{\lambda} = \left(\lambda, \boldsymbol{\lambda}^v, \lambda^W, \boldsymbol{\lambda}^W\right) : \ \boldsymbol{\lambda} \in \mathbb{R}^8, \lambda^W > 0 \quad \text{and} \quad v_\infty |\boldsymbol{\lambda}^W| < \lambda^W\right\},$$

where $v_\infty = 1/\sqrt{2\alpha m^*}$ is the asymptotic value of $v(\mathcal{E})$.
Again, the additional moments G^{ij} do not introduce further restrictions on the domain of definition Λ because

$$\frac{1}{\hbar}\frac{\partial}{\partial k_j}(\mathcal{E} v_i) = v_i v_j + \frac{\mathcal{E}}{\hbar}\frac{\partial v_i}{\partial k_j}$$

is integrable together with the exponential densities $f_{\boldsymbol{\lambda}}$.
In the next considerations we limit ourselves to the case when $\boldsymbol{u}$ and $\boldsymbol{S}$ are collinear. This is certainly verified for 1-D problems.

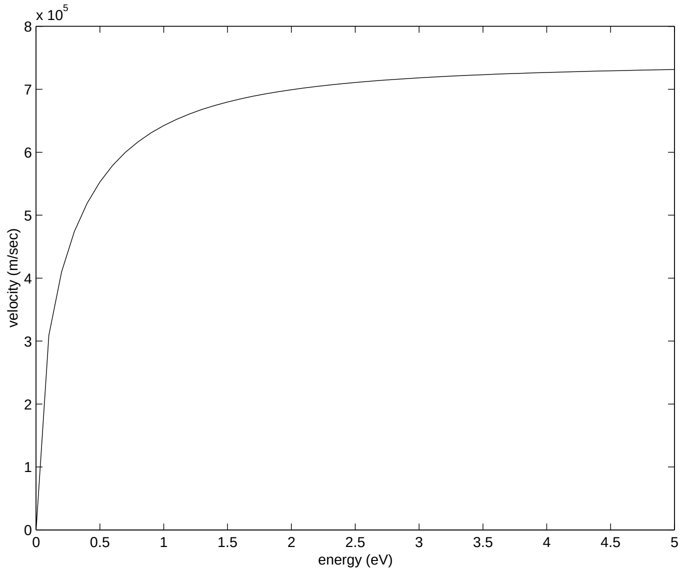

Figure 2: modulus of $\boldsymbol{v}$ as function of $\mathcal{E}$. Note that it is an increasing function with asymptotic value $v_\infty = 1/\sqrt{2\alpha m^*}$ which is about 7.4 $\times 10^5$ m/sec in Silicon.

By choosing a reference frame $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ adapted to $\boldsymbol{u}$, one can write $\boldsymbol{u} = u_3\mathbf{e}_3$, $\boldsymbol{S} = S_3\mathbf{e}_3$, $\boldsymbol{\lambda}^v = \lambda_3^v\mathbf{e}_3$ and $\boldsymbol{\lambda}^W = \lambda_3^W\mathbf{e}_3$.
The explicit relation between the Lagrange multipliers and the macroscopic variables are given by

$$
\begin{aligned}
n &= 4\sqrt{2}\pi\frac{(m^*)^{3/2}}{\hbar^3}\, e^{-\lambda} d_0, && (43)\\
u_3 &= \frac{1}{d_0}\int_0^\infty v(\mathcal{E}) e^{-\lambda^W\mathcal{E}}\sqrt{\mathcal{E}(1+\alpha\mathcal{E})}(1+2\alpha\mathcal{E})\left[\frac{\sinh y}{y^2} - \frac{\cosh y}{y}\right] d\mathcal{E}, && (44)\\
W &= \frac{1}{d_0}\int_0^\infty \mathcal{E} e^{-\lambda^W\mathcal{E}}\sqrt{\mathcal{E}(1+\alpha\mathcal{E})}(1+2\alpha\mathcal{E})\,\frac{\sinh y}{y}\, d\mathcal{E}, && (45)\\
S_3 &= \frac{1}{d_0}\int_0^\infty \mathcal{E} v(\mathcal{E}) e^{-\lambda^W\mathcal{E}}\sqrt{\mathcal{E}(1+\alpha\mathcal{E})}(1+2\alpha\mathcal{E})\left[\frac{\sinh y}{y^2} - \frac{\cosh y}{y}\right] d\mathcal{E}, && (46)
\end{aligned}
$$

with

$$
\begin{aligned}
y &= \left(\lambda_3^v + \mathcal{E}\lambda_3^W\right)\, v(\mathcal{E}), && (47)\\
d_0 &= \int_0^\infty e^{-\lambda^W\mathcal{E}}\sqrt{\mathcal{E}(1+\alpha\mathcal{E})}(1+2\alpha\mathcal{E})\,\frac{\sinh y}{y}\, d\mathcal{E} && (48)
\end{aligned}
$$

Again there is not limitation on n. Therefore it is relevant only to study the dependence of u_3, W and S_3 on λ_3^v, λ^W and λ_3^W. In order to have an estimate about the moment cone $\mathcal{M}$ we have numerically evaluated the relations (44)-(46).
In figure 3 there are plotted some cross-sections of the moment cone obtained as images of cross-sections of Λ under the mapping $(\lambda_3^v, \lambda^W, \lambda_3^W) \mapsto (u_3, W, S_3)$ defined by the relations (44)-(46). $\lambda_3^v/\sqrt{m^*}$ is expressed in $1/\sqrt{\text{eV}}$, λ^W in 1/eV and $\lambda_3^W/\sqrt{m^*}$ in $1/\sqrt{\text{eV}^3}$.

We have considered the images of the cuts

$$\left\{\lambda_3^W/\sqrt{m^*}=0,\ (\lambda_3^v/\sqrt{m^*},\lambda^W)\in[-15,15]\times[5,74]\right\},$$
$$\left\{\lambda_3^W/\sqrt{m^*}=10,\ (\lambda_3^v/\sqrt{m^*},\lambda^W)\in[-15,15]\times[20,39]\right\},$$
$$\left\{\lambda_3^W/\sqrt{m^*}=15,\ (\lambda_3^v/\sqrt{m^*},\lambda^W)\in[-15,15]\times[20,39]\right\},$$
$$\left\{\lambda_3^W/\sqrt{m^*}=-10,\ (\lambda_3^v/\sqrt{m^*},\lambda^W)\in[-15,15]\times[20,39]\right\}.$$

The figure 3 shows that also in the 8-moment model the moment cone is sufficiently wide to enclose the relevant physical region of density, velocity, energy and energy-flux.
At last we have numerically inverted the relations (44)-(46) for given values of the moments. n is set equal to one in arbitrary units without loss of generality because it influences only the value of λ. The corresponding distributions, normalized as $f_{ME}/(2\sqrt{2}\pi\frac{(m^*)^{3/2}}{\hbar^3})$, have been plotted versus the energy and $\mathcal{E}$ and $\cos\theta$ since in the one dimensional case there is symmetry with respect to the angle ϕ.
In figure 4 we compare the results for two sets of values of moments

$$W=W_0,\ u=10^3\text{m/sec}, S=0 \text{ and } \ W=W_0,\ u=10^4\text{m/sec}, S=0,$$

where $W_0=0.039$ eV is the equilibrium energy at the room temperature of 300 ^{o} K. For the multipliers $\lambda_3^v,\lambda^W,\lambda_3^W$ one finds (in the same units used above)

$$\lambda_3^v/\sqrt{m^*}=-0.1982, \lambda^W=39.6472, \lambda_3^W/\sqrt{m^*}=2.2545$$

in the first case and

$$\lambda_3^v/\sqrt{m^*}=-1.9139, \lambda^W=39.9371, \lambda_3^W/\sqrt{m^*}=21.2887$$

in the other case. One can observe that the maximum of the distribution increases with the velocity.
In figure 5 we compare the results for other values of moments

$$W=0.2\text{eV}, u=10^3\text{m/sec}, S=10^4\text{m eV/sec and } \ W=0.2\text{eV}, u=10^5\text{m/sec}, S=10^4\text{m eV/sec},$$

One gets the multipliers

$$\lambda_3^v/\sqrt{m^*}=1.1109, \lambda^W=8.8073, \lambda_3^W/\sqrt{m^*}=-3.8693$$

in the first case and

$$\lambda_3^v/\sqrt{m^*}=-3.5707, \lambda^W=8.7404, \lambda_3^W/\sqrt{m^*}=6.7119$$

in the other case. By increasing S, $\lambda_3^W/\sqrt{m^*}$ tends to the critical value λ^W and, as consequence, the maximum of f_{ME} increases. The same happens for the last set of moment

$$W=0.3\text{eV}, u=10^5\text{m/sec}, S=10^4\text{m eV/sec and } \ W=0.35\text{eV}, u=10^5\text{m/sec}, S=10^4\text{m eV/sec},$$

for which we have the multipliers

$$\lambda_3^v/\sqrt{m^*}=-2.9952, \lambda^W=6.0804, \lambda_3^W/\sqrt{m^*}=4.2214$$

and

$$\lambda_3^v/\sqrt{m^*}=-2.8077, \lambda^W=5.3011, \lambda_3^W/\sqrt{m^*}=3.5276.$$

The corresponding distributions are plotted in figure 6.

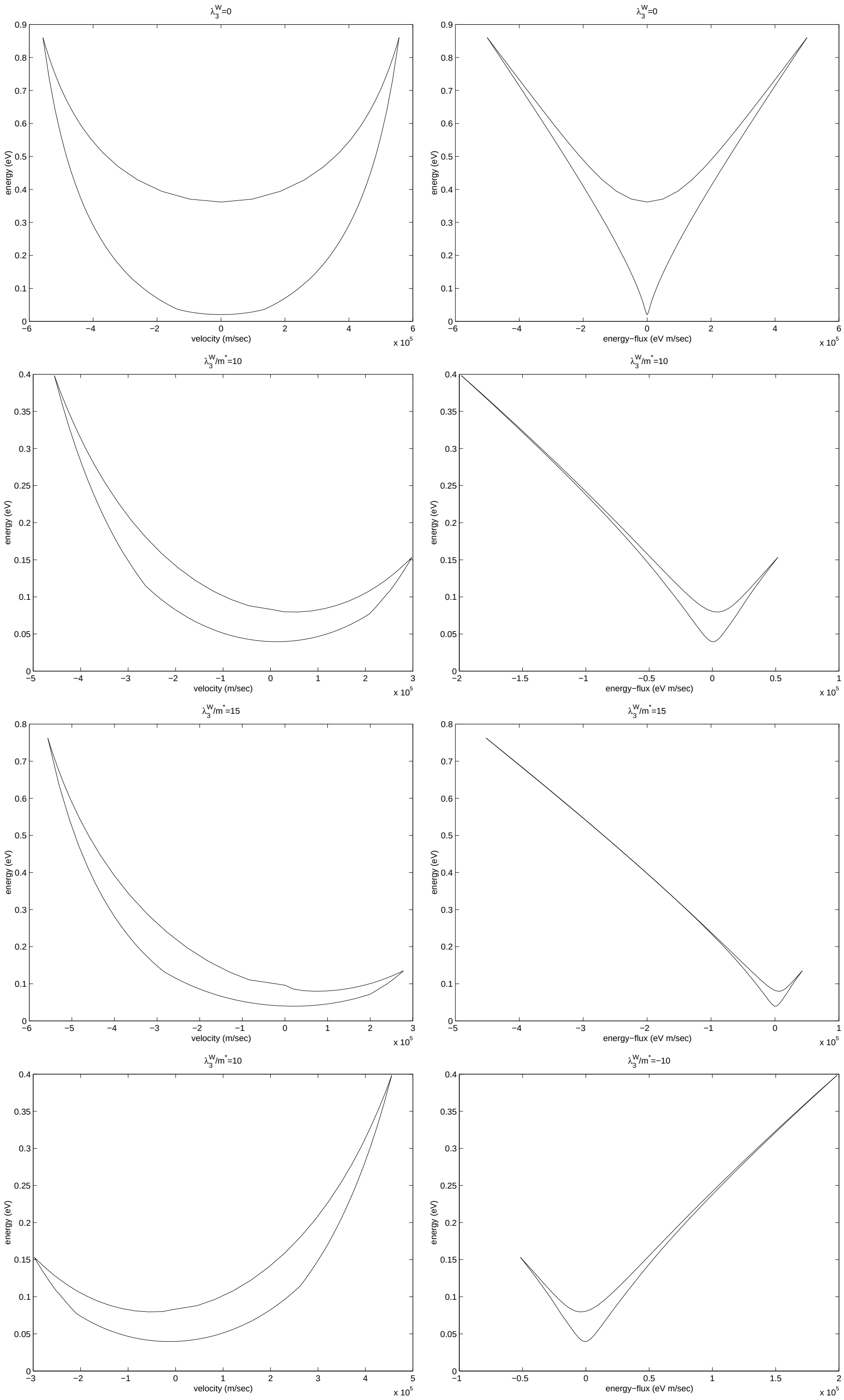

Figure 3: Some cross-sections of the moment cone obtained as images of Λ under the mapping $(\lambda_3^v, \lambda^W, \lambda_3^W) \mapsto (u_3, W, S_3)$ defined by the relations (44)-(46).

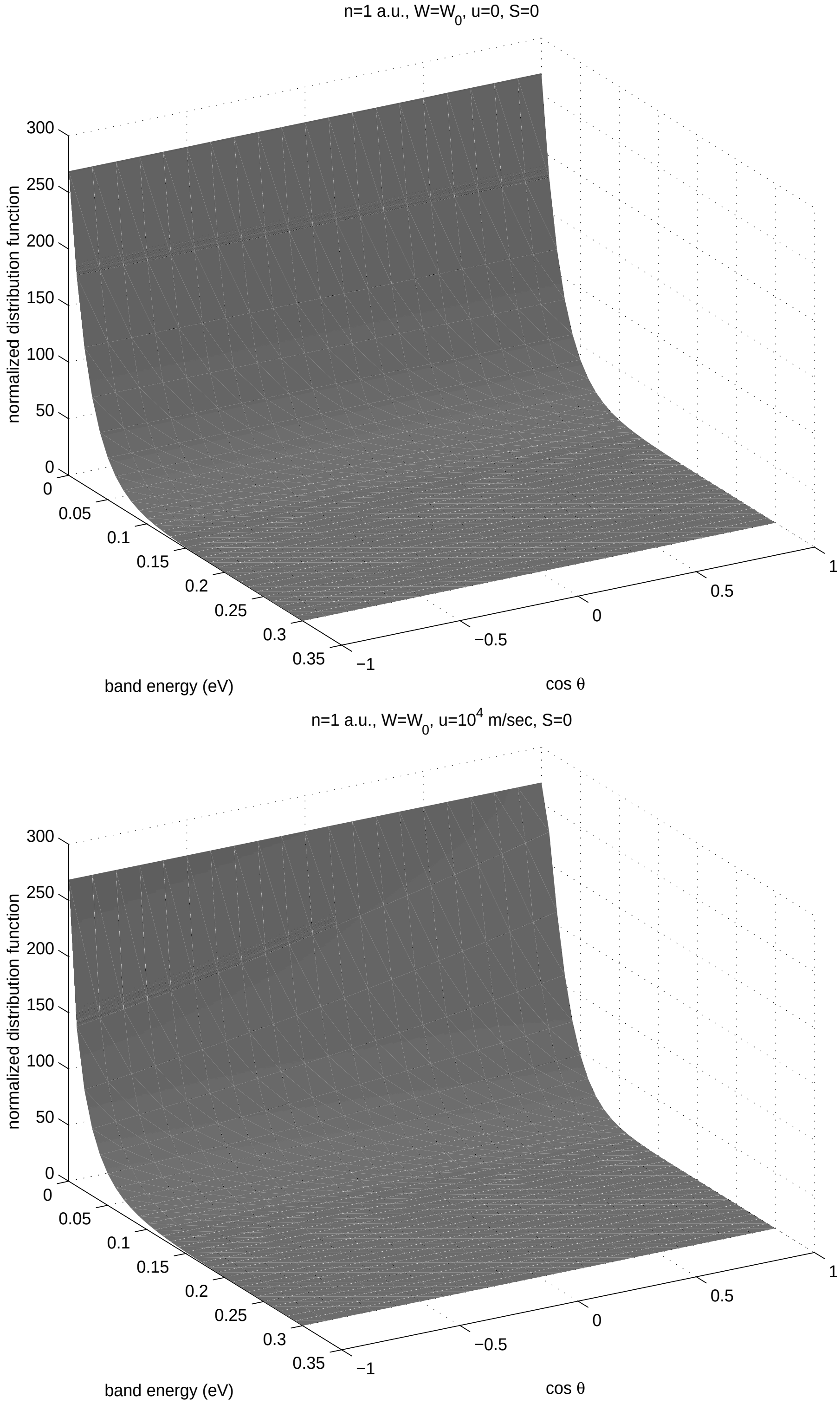

Figure 4: $f_{ME}/(2\sqrt{2}\pi\frac{(m^*)^{3/2}}{\hbar^3})$ for the two set of moments $n = 1a.u., W = W_0$, $u = 10^3\mathrm{m/sec}, S = 0$ and $n = 1a.u., W = W_0$, $u = 10^4\mathrm{m/sec}, S = 0$.

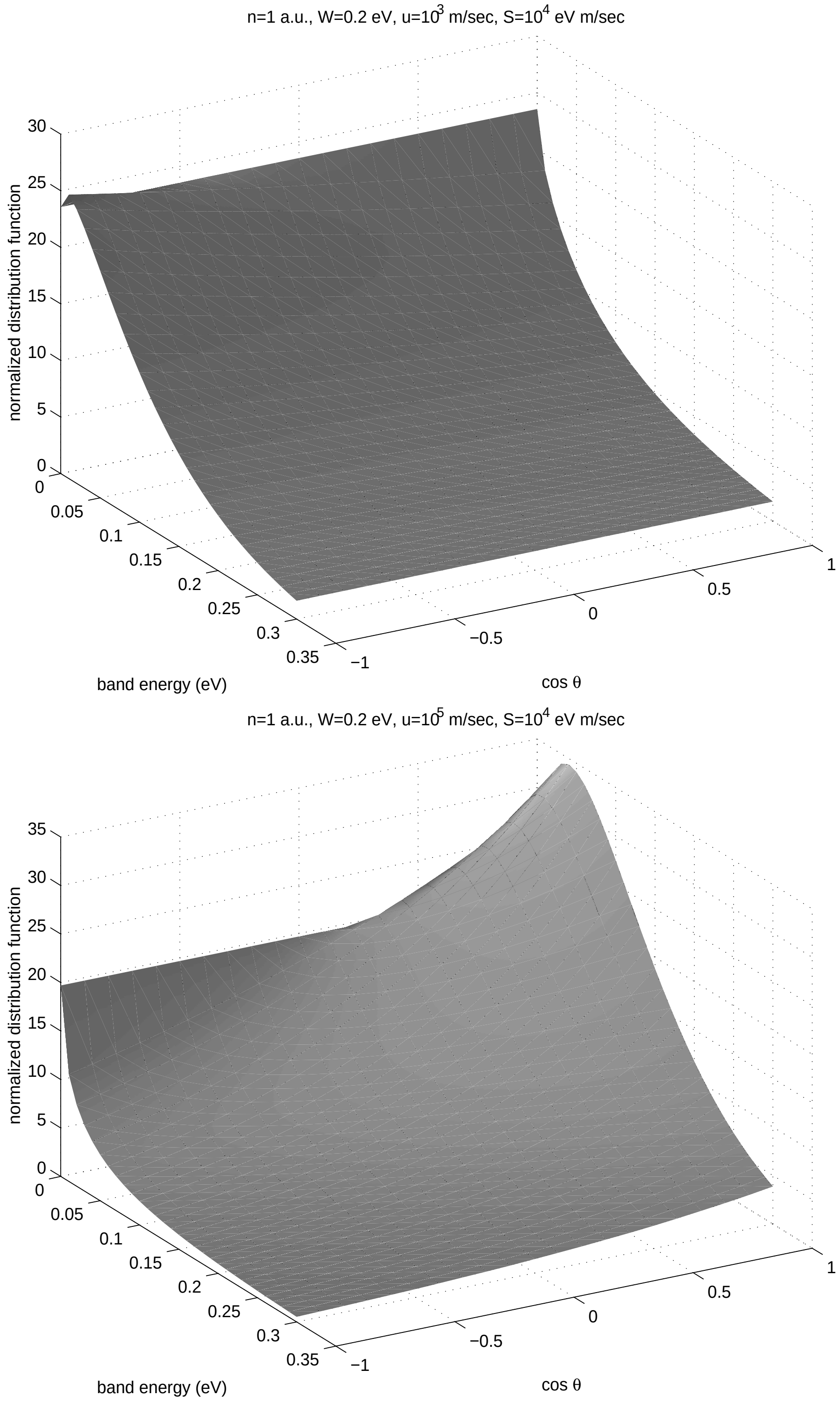

Figure 5: $f_{ME}/(2\sqrt{2}\pi\frac{(m^*)^{3/2}}{\hbar^3})$ for the two set of moments $n = 1a.u., W = 0.2\text{eV}, u = 10^3\text{m/sec}, S = 10^4\text{m eV/sec}$ and $n = 1a.u., W = 0.2\text{eV}, u = 10^5\text{m/sec}, S = 10^4\text{m eV/sec}$.

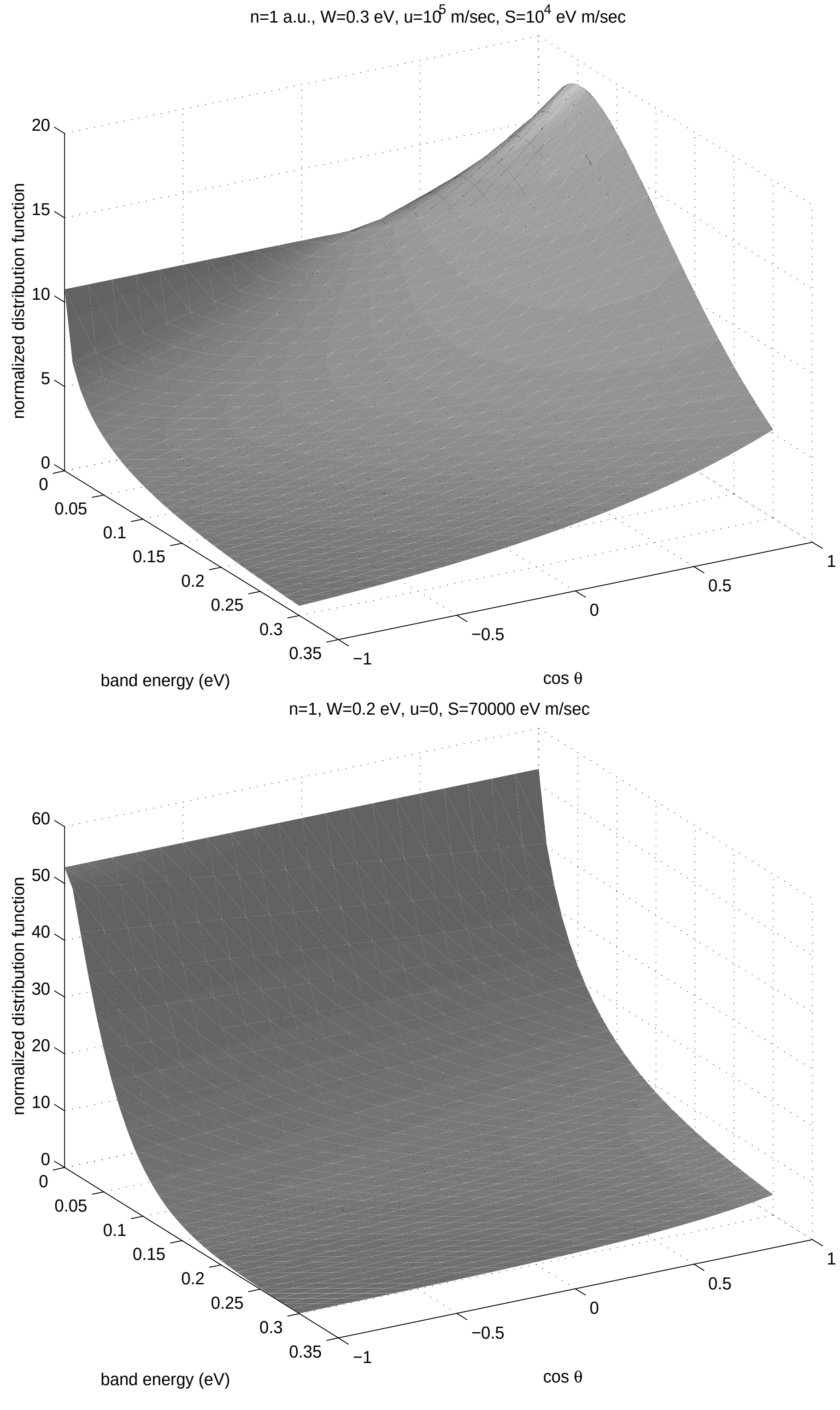

Figure 6: $f_{ME}/(2\sqrt{2}\pi\frac{(m^*)^{3/2}}{\hbar^3})$ for the two set of moments $n = 1a.u., W = 0.3\text{eV}, u = 10^5\text{m/sec}, S = 10^4\text{m eV/sec}$ and $n = 1a.u., W = 0.35\text{eV}, u = 10^5\text{m/sec}, S = 10^4\text{m eV/sec}$.

5 Acknowledgments

The author V.R. acknowledges the financial support by M.I.U.R. COFIN 2002 *Problemi Matematici delle teorie cinetiche*), P.R.A. (ex 60 %), CNR (grant n. 00.00128.ST74) and TMR project *Hyperbolic and kinetic equations: asymptotics, numerics and applications* grant HPRN-CT-2002-00289.

References

[1] A.M. Anile, G. Mascali, V.Romano. Recent developments in hydrodynamical modeling of semiconductors (2003) 1:54 in Mathematical Problems in Semiconductor Physics, Lecture Notes in Mathematics 1832, Springer (2003).

[2] C. E. Shannon, , Bell System Tech. J. 27: 379, 623, also reprinted in C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, 1949.

[3] E. T. Jaynes, *Information theory and Statistical Mechanics*, Phys. Rev. 106: 620–630, 1957.

[4] W. Dreyer. Maximization of the entropy in non–equilibrium. *J. Phys. A: Math. Gen.*, 20:6505–6517, 1987.

[5] C. D. Levermore. Moment closure hierarchies for kinetic theories. *J. Stat. Phys.*, 83:1021–1065, 1996.

[6] M. Junk. Domain of definition of Levermore's five-moment system. *J. Stat. Phys.*, 93:1143–1167, 1998.

[7] M. Junk. Maximum entropy for reduced moment problems. *Math. Models Methods Appl. Sci.*, 10:1001–1025, 2000.

[8] M. Junk. *Moment problems in kinetic theory.* Universität Kaiserslautern, Habilitationsschrift, 2001.

[9] P. Markowich, C. Ringhofer, and C. Schmeiser. *Semiconductor equations.* Springer, Wien, 1990.

[10] N. C. Ashcroft and N. D. Mermin. *Solid State Physics*, Philadelphia, Holt-Sounders (1976).

[11] C. Jacoboni and L. Reggiani, *The Monte Carlo method for the solution of charge transport in semiconductors with application to covalent materials*, Review Modern Phys. 55 (1983), pp. 645–705.

[12] C. Jacoboni and P. Lugli, *The Monte Carlo Method for Semiconductor Device Simulation*, Wien - New York, Springer-Verlag (1989).

[13] T. Ruggeri, *Galilean invariance and entropy principle for systems of balance laws. The structure of extended thermodynamics*, Cont. Mech. Thermodyn., 1 (1989) pp.3–20.

[14] A. M. Anile and V. Romano , *Non parabolic band transport in semiconductors: closure of the moment equations*, Cont. Mech. Thermodyn., 11 (1999) pp. 307–325.

[15] B. C. Eu. A modified moment method and irreversible thermodynamics. *J. Chem. Phys.*, 73:2958–2969, 1980.

[16] A. N. Gorban and I. V. Karlin. Method of invariant manifolds and the regularization of acoustic spectra. *Transp. Theory Stat. Phys.*, 23:559–632, 1994.

[17] I. Müller and T. Ruggeri. *Rational Extended Thermodynamics.* Springer, New York, 1998.

[18] D. Jou, J. Casas-Vazquez and G. Lebon, *Extended irreversible thermodynamics*, Springer-Verlag, Berlin, 1993.

[19] A. M. Anile, V. Romano and G. Russo, *Extended hydrodynamical model of carrier transport in semiconductors*, SIAM J. Appl. Math., 61 (2000), pp. 74–101.

[20] V. Romano, *Non parabolic band transport in semiconductors: closure of the production terms in the moment equations*, Cont. Mech. Thermodyn., 12 (2000) pp. 31–51.

[21] V. Romano, *Non parabolic band hydrodynamical model of silicon semiconductors and simulation of electron devices*, Math. Meth. Appl. Sci. 24 (2001) pp. 439–471.

[22] G. Mascali and V. Romano, *Hydrodynamical model of charge transport in GaAs based on the maximum entropy principle*, Cont. Meh. Thermodyn., 14 (2002) pp. 405–423.

[23] W. Dreyer, M. Junk, and M. Kunik. On the approximation of the Fokker-Planck equation by moment systems. *Nonlinearity*, 14:881–906, 2001.

[24] M. Junk and A. Unterreiter. Maximum entropy moment systems and Galilean invariance. *Continuum Mech. Thermodyn.*, 14:563–576, 2002.

[25] A. Majorana. Conservation laws from the Boltzmann equation describing electron-phonon interactions in semiconductors. *Transp. Theory Stat. Phys.*, 22: 849-859, 1993.

[26] A. Majorana. Equilibrium solutions of the non-linear Boltzmann equation for an electron gas in a semiconductors. *Il Nuovo Cimento*, 108B: 871–877, 1993.

[27] V. Romano. Maximum entropy principle for electron transport in semiconductors. Proceedings InternationalConference WASCOM 99, Vulcano (Eolie Isles) Italy, June 1999.

[28] I. Csiszár. I-divergence geometry of probabaility distributions and minimization problems. *Ann. of Prob.*, 3:146–158, 1975.

[29] A. S. Lewis. Consistency of moment systems. *Can. J. Math.*, 47:995–1006, 1995.